

Data Elicitation for Continuous Awareness of Team Climate Characteristics to Improve Organizations' Creativity

Jan Nierhoff
University of Bochum
jan.nierhoff@rub.de

Thomas Herrmann
University of Bochum
thomas.herrmann@rub.de

Abstract

The creativeness of a company's employees depends on the characteristics of working climates, e.g. autonomy or appropriate workload. Tools for their assessment exist, but the frequency of their application is too low to detect the relevant dynamics which characterize the varying challenges of agile and learning organizations. The evaluation of a first prototype to monitor these dynamics by frequently repeating a common online employee survey revealed relevant features to overcome a lack of acceptance of answering the same question items in repetition.

Three variables were identified which influence the acceptance of a repeated question: The time since it had last been answered, the user's current willingness to participate and the user's situation. Based on these variables, a new prototype offers users more self-determination in their rate of participation, allows for assigning dynamic repetition rates to every question item, and exploits context information to optimize the prompting of users.

1. Introduction

This paper focusses on methods of investigating organizations' working climates as a crucial precondition for creativity in teams. Climate is not the only factor, but a crucial one that influences the creativity of employees, and it is widely acknowledged as a key parameter in theories that deal with organizational creativity [1]–[3]. Perception of creativity in a working climate involves a great deal of inertia; if the climate worsens and thereby affects the employees' perception of it negatively, it takes a long time for a working climate to recover to its initial state, even if the original conditions are restored [4]. This makes the timely detection of a work climate's dynamics in relation to creativity a critical task [5].

When assessing the climate, the aggregated perception of working environment characteristics by the employees is more useful than an objective description

by third parties [6], [7]. Thus, employee surveys are the primary approach for measuring working climate [8]. However, such surveys are designed to measure a certain state of an organization at a certain point in time within longer periods, such as once a year.

Thinking of modern agile work scenarios in learning organizations, where employees regularly switch between different projects and teams, leads to the assumption that the working climate and the perception of it potentially change at a higher frequency. This is due to more variations of conditions than conventional employee surveys can detect. These dynamics usually influence other work conditions, such as stress or the extent of surveillance, both of which are highly correlated to the degree of possible creativity [9], [10]. Stress is a phenomenon for which dynamic variations over time can be observed (see 3.2). Since we comply with approaches which emphasize the possibility of promoting and facilitating creativity [11], [12], we emphasize the relevance of timely feedback on dynamic changes or interventions and their effect on creativity.

Furthermore, common surveys ask about aggregated subjective perceptions and about general attitudes instead of hard facts, but the answers are probably influenced by the participants' situation and mood at the time of the survey [13]. To reduce that influence we strive for a survey system which enables participants to distribute answers over multiple points in time.

Also, the questions in common survey tools are formulated in a way that is not conducive to continuous short-term repetition. They are more focused on people's attitudes that have emerged over a longer period instead of asking how people have perceived events within a specific, recent time span.

With project "KreativBarometer" [14] we intend to address these shortcomings by creating a survey system (see Section 3) that is based on the common employee survey design, but in addition

- is continuously employed with a repetition rates of every few weeks, publishing fresh results after every repetition (see Figure 1),

- enables participants to give answers scattered over time
- rephrases question items from established creativity assessment tools by having a specific temporal focus, e.g. “Currently, I feel accepted and understood by my peers.”.

Based on the computed results, employees can compare their own views with the average results of their team and the whole company. The repetitive character of the survey also allows for monitoring the development of the climate with new data sets every few weeks. The central research question is: What are a survey tool’s crucial features which unobtrusively remind and prompt selected employees to continuously answer questions that deal with creativity related conditions of their work, and encourages them to distribute their answers over time. The data collection rate should be sufficient to produce results that can be mapped to specific situations or events related to creativity.

We first implemented a solution as to how to repeat conventional surveys and have evaluated (Section 3) how this approach affected the participation rates. Several deficits and acceptance barriers became apparent. Based on these results we have derived a more sophisticated design concept (see Section 4), which deals with the trade-off between allowing the participants more self-determination and maintaining a sufficient repetition rate. Section 5 summarizes the main characteristics of the redesign and points to further research needed in relation to creativity.

2. Related Research: Employee surveys, task interruptions, and context awareness

Employee Surveys – established, but insufficient sampling rates. The literature on employee surveys reveals that they are well established on the market and a de-facto standard for participatory human resource-, organizational- and corporate development (Over 75% of companies that participated in studies in US-America [15] and the German-speaking area of the EU [16] do employee surveys.). They are often implemented repetitively as longitudinal studies but feature quite low repetition rates of one to two years, and they demand on average an organizational overhead of 8 months [16]. Conducting employee surveys online probably reduces that overhead. Nonetheless, employee surveys are conducted as singular events, gathering the aggregated mood of a group of people at a certain point in time. The described characteristics of common surveys are not suitable for modern, agile working scenarios in which employees may switch teams and projects/tasks at high rates. While agile con-

cepts, like heterogeneous, dynamic teams or short development cycles, have positive effects on the innovation power of companies, they are also likely to affect factors that influence individual and team creativity, such as the size of the workload.

The literature emphasizes acceptance by all stakeholders as a critical factor of influence on the success of an employee survey. Skepticism and refusal can lead to insufficient participation. Unconsciously or knowingly, the participants may give inappropriate answers or lack the willingness to act upon the results of the survey [17]. Securing participants’ anonymity via privacy and data security mechanisms is mandatory to build the needed acceptance of the tool on the employee side [18], [19].

Task interruption. Related research deals with understanding how much of an interruption the answering of question items implies and how this interruption may influence the survey. Research on task interruptions endorses acceptance as a critical factor of success. Besides the risk of bothering users too much and causing that they discontinue participating in the survey, there is also the risk of negatively influencing the answering behavior, because annoyance created by an interrupting tool can lead to prejudice against the content it delivers [20]–[22].

Because the survey takes place in the context of work, users may be required to answer items concurrently to other tasks, which implies a distraction of some degree. A typical team climate question asks, for example: “Within the last 2 weeks, I felt accepted and understood by my teammates”. Answering requires memory retrievals which means an overriding of a former problem representation the user has in mind (i.e. of the interrupted task) and drastically hampers task resumption. The user is thereby disrupted and potentially annoyed. This applies even if the time needed for an answer is relatively short [23], [24]. As a consequence, a tool for continuous surveys should only inform the user about the availability of a question instead of making it appear immediately. This strategy decreases the resumption lag when returning to the previous task [25]. Using the standard notification mechanisms offered by the operating systems can also help to be less obtrusive since users are already accustomed to them.

Research suggests avoiding the disruption of users while they are working on so-called fine-level tasks (e.g. composing an email or reading a webpage). Using changes between coarse-level tasks (like switching from work on project A to doing standard paperwork) to notify users decreases annoyance [26].

In summary, the continuous collection of opinions creates a dilemma: At the one hand it is reasonable to ask for work-related experience directly in the context

of work. On the other hand, the kind of questions that are asked result in an interruption of a primary work task. One approach to solve this is to detect task switches to exploit these for prompting employees to answer a question.

Context Awareness – Computer Systems Cannot Assess Cognitive Processes. Context awareness is the concept of discovering and taking advantage of contextual information to anticipate users' needs and act accordingly. Research on notification systems brought up context aware systems that try to assess the user's attentional focus to minimize disturbance (e.g. Microsoft's "Attentional User Interface Project" [27]). Further investigations into this class of context awareness reveal that they work satisfactorily in controlled, confined environments but not necessarily in real-world scenarios.

While context aware features are more and more common (especially the use of user location or examples of contextual assistants like Apple's Siri, Google's Google Now or Microsoft's Cortana), the available exploitation possibilities are still limited. Existing systems are based on external, physical contexts [28] and do not make assumptions about cognitive or emotional domains, like happiness or stress. Mastering a generic task like knowing when a user is willing to answer a question is not realizable in real-life scenarios. Nevertheless, the survey tool that is presented in Section 4 uses context information in an attempt to optimize the point in time to prompt a user for the sake of maximizing continuous user acceptance. Because detecting task switches is only partly achievable under real work conditions, we additionally look into ways to unobtrusively notify and remind users about the request to answer survey items.

Interruptions, particularly at inappropriate points in time, can be an interference especially in the context of creativity, since they interrupt the flow which is characterized as an essential mode of creative activity [29]. Consequently, the survey has to be designed in a way that it gives timely feedback to changes and effects of interventions which influence creativity without disturbing the creative flow itself.

3. Conducting a Conventional Survey with High Repetition Rates

The prototypic implementation of a survey system was based on a conventional employee survey design (especially the questionnaire design and result computation). Repetition was realized by starting one survey iteration after another. The procedure of data collection was also unconventional, since the participants could answer a single set or small sets of items at different points in time, instead of in one sole session.

3.1. Prototypic Implementation

Based on the presented literature research and the goal to use surveys repetitively to monitor dynamic changes in team culture, an initial series of prototypes was implemented with the following characteristics: Participants were given specific timeframes (usually between one and four weeks) in which they could answer the questionnaire. The questionnaire consisted of 25 items that were selected from established tools for creativity climate assessment (Keys – Assessing the Climate for Creativity [1], Team Climate Inventory [30], Creativity and Work Environment [31]). We rephrased the items to focus on the last few weeks prior to the survey, instead of addressing a general, time-unspecific opinion. Between the consecutive survey iterations, the results were computed analogously to a conventional survey (see Figure 1).

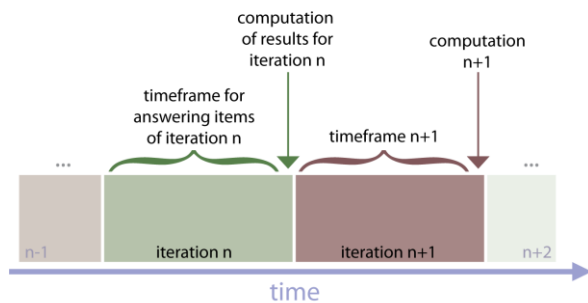


Figure 1. Modus operandi of the survey system.

To answer the questions, the participants were offered a browser-based or a desktop-application client (see Figure 2). While the browser-based client provided an easier technical integration at the test sites, the desktop-application featured prompting functionalities on a more sophisticated level and minimized the number of mouse clicks that were needed to answer questions. Both clients used the approach of consecutively presenting one single item at a certain point in time to the participants. Users of the desktop client were free to choose from and configure the following functionalities or rules:

- Deadline warning – a balloon tip that warned the users x days before the end of a survey iteration in case they had not answered all items. This functionality aimed at maintaining sufficient participation.
- Return to workstation prompt – a balloon tip or an item that popped up if the users interacted with the computer after y minutes of idleness. This reminder aimed at prompting the user when s/he was switching between coarse-level tasks.

- Auto close – the application’s GUI closed automatically after z items had been consecutively answered. This functionality aimed at promoting the distribution of answers over a certain time span.

- Daily item limit – initially we limited the number of answers that could be given per day to three. This was supposed to encourage a distribution of answers over a certain period. This restriction was deactivated during the evaluation period, as several participants complained about it.

The prototypes also complied with ethical factors which literature research on employee surveys identified as critical (see 2). The participants used pseudonymous accounts, which were not linked to any personal data except their team affiliation. Results could only be accessed by the individuals who produced them. Data security was granted by SSL end-to-end encryption and secured servers and databases. Furthermore, a socio-technical process was organized that provided transparency about how, when and by whom the results were put to use and informed participants about privacy and data security. Additionally, we pointed out to our industrial partners that taking part in the survey should be voluntary.

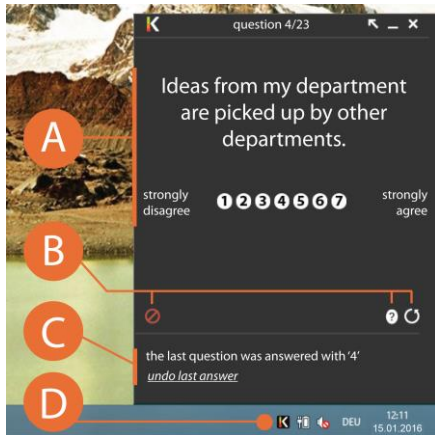


Figure 2. Desktop client: A) item text & answer buttons B) decline or switch the current item C) undo last answer D) representation in notification area.

3.2. Evaluation of the Conventional Approach

To do a formative evaluation of the initial prototypes and to allow for continuous, but incremental improvement, field tests were run at several sites (sites 1 to 10 in Table 1). These tests were included into the overall goal of giving feedback as to whether and how a company could improve its creativity climate.

In addition, a summative evaluation was conducted with a team of 12 office workers (see site 11 in Table 1 and Figure 3). Six survey iterations were completed within this last experiment. The second iteration lasted

one week and the others three weeks. The participants used the desktop-application to answer a questionnaire of 26 items per iteration. We applied a mixed method approach, analyzing the usage data of all participants and interviewing seven participants.

Usage statistics showed that the prototype encouraged and enabled clients to participate sufficiently over the time of the experiment. Figure 3 illustrates that the participation during the 16-week long experiment fluctuated, but did not decrease. One may argue that this was due to the experimental setup of the usage scenario and may not hold true for longer periods of time or even an open-ended implementation. Nonetheless, an observation at Test Site 9, at which a team autonomously continued using the tool at a satisfactory rate for 17 additional survey iterations after the experiment, implies that this initial approach was suitable to maintain a satisfactory participation rate over time, if the appropriate socio-technical circumstances were given.

Test site	Users	Teams	Iterations in total	Answers given in total
1	7	1	10	1876
2	14	1	16	4755
3	54	5	17	5740
4	87	5	25	31649
5	15	1	31	4293
6	43	6	11	3895
7	9	1	10	1151
8	46	2	94	5519
9	62	7	25	5009
10	128	12	7	4301
11	12	1	6	1554
all	477	42		69742

Table 1. Statistics of the field tests.

On the downside, quantitative data revealed that the usage of the prototype did not result in the desired behavior of answering small sets of items distributed over time. One participant showed the desired behavior, while ten answered all 26 items of an iteration within one coherent session. One participant showed the desired behavior in the first iteration, but changed to “bulk-answering” afterwards, her answering behavior is depicted in Figure 4.

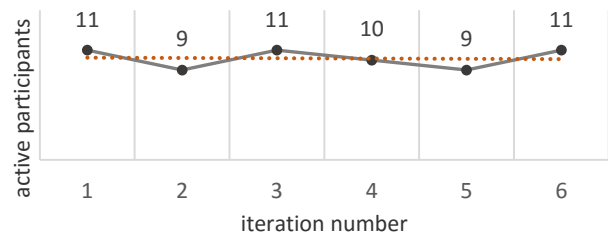


Figure 3. Sufficiently active users during six iterations.

While distributing the answers over time was not our primary concern, the explanations which participants gave in interviews for this behavior implied an acceptance problem. Participants stated that answering the survey felt like “doing homework,” and a favored attitude of dealing with it was “getting it done as fast as possible” so that the system “leaves you be for some weeks”. A negative attitude towards the survey was something that we definitely needed to avoid, as literature research on employee surveys and task interruptions suggest. More qualitative data on the surveys ‘acceptance problem’ was collected:

The interviews revealed that there were **intra- and inter-personal differences** in how many items per time users were willing to answer. The number of answers that users were willing to answer averaged at 13.7 per week, with a standard deviation of 8.4 and with 5 and 25 items per week at the extremes. On request, five interviewees responded that their statement was probably due to changes in contextual factors (e.g. fewer answers in stressful times or more answers to let off steam when upset). Another outcome of the interviews was that **the interest in answering items repeatedly depended on the type of question**. The questions we used stemmed from surveys that ask for the general opinion of participants and focus on topics that may be subject to different temporal dynamics. For example, participants may be willing to report about their current stress level on a weekly basis, but find no purpose in assessing their ‘work-place’s functionality’ as frequently, since it has more stable conditions. Also notable was the participants’ **little tolerance for rules** or restrictions regarding the answering process. As mentioned before, the daily item limit was deactivated already during the project’s tests as it earned harsh criticism. When asked about that refusal, users stated that the restriction made them feel patronized and they perceived it as an artificial impediment.

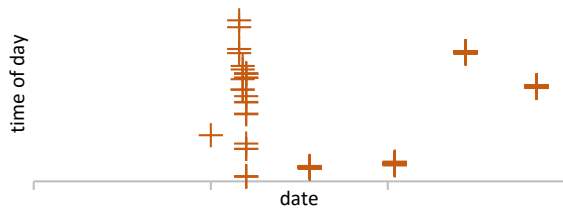


Figure 4. Points of time at which items have been answered by one specific participant. The thicker a mark’s crossbar, the more items were answered.

In summary, the evaluation resulted in the following challenge: Create a survey system that maximizes user acceptance. Do this by respecting inter- and intrapersonal dynamics in the willingness to answer items and by assigning reasonable frequencies at

which specific items should be answered. This and answering behavior distributed over time should be promoted without enforcement by rules or superiors that may make the users feel patronized.

In addition to insights about the continuous data-collection, the field tests allowed us to get first insights into creativity climate dynamics. While not all test sites and examined dimensions showed significant changes within short periods of time, some dynamics were noticeable and strengthen our argumentation for continuous monitoring of the creativity climate. E.g. Figure 5 shows the dimension ‘stress’ at Test Site 3.

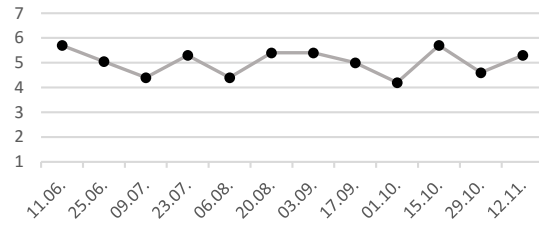


Figure 5. Climate dynamics of dimension ‘stress’

4. Evaluation-based Redesign: Self-Determined Participation and a Dynamic Questionnaire

This section shows the evaluation-based redesign of the survey system: Section 4.1 introduces the concept of a dynamic questionnaire which promotes distributing answering over a longer time period by indicating a subset of the questionnaire’s items as ‘relevant’ at a specific point in time. The decision as to whether an item is relevant is based on the time that has passed since it has last been answered and the user’s estimated current willingness and interest in answering it. The estimation about the user’s willingness is based on direct feedback from the user on the one hand, and inferences from contextual information on the other hand. The open context awareness framework is presented in detail in Section 4.2. Furthermore, the redesign was developed to be applied onto smartphones to allow for mobile usage, since the evaluation revealed the necessity for more flexibility for the users.

4.1. A Dynamic Questionnaire for Continuous Data Collection

This section introduces a survey questionnaire that dynamically adapts by defining only a subset of its items as relevant at specific points in time. The users’ attention is directed towards this subset.

To determine the relevancy of a question item, three input variables are taken into account: The item's frequency, the user's current willingness to participate and contextual information to (dis-)favor the current situation for the sake of minimizing disruption.

To integrate these inputs in one system, a threshold concept is applied. Each item has a current *urgency value* that is based on its frequency and each user has a dynamic *acceptance threshold* that represents his or her current willingness to answer survey items. If an item's urgency exceeds the acceptance threshold, it is determined to be *relevant* at that point in time. Contextual information is taken into account by letting it influence the user's acceptance threshold. If, for example, a favorable situation for answering items is detected, a negative offset is applied to the threshold. This may make items relevant, if they exceed the temporarily decreased threshold. After the favorable situation passes, the threshold returns to its original value. Figure 6 visually represents how a user's acceptance threshold 'hides' items with a lower urgency level from the user. Hidden items can become relevant when their own urgency increases with time or when the acceptance threshold decreases.

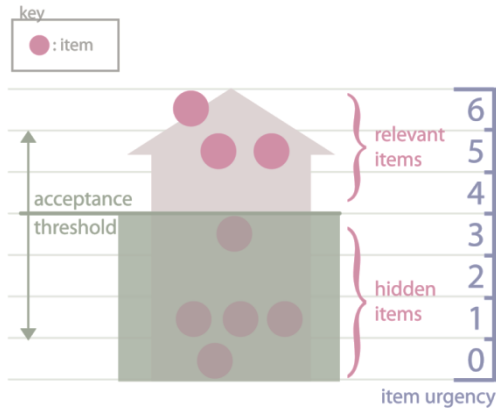


Figure 6. The items' urgency levels rise with time. The user's acceptance threshold determines which items are relevant for being answered and which are hidden.

Definition of item relevance: An item is relevant for a specific user at a specific point in time, if its urgency value exceeds the user's acceptance threshold at that point in time.

Current urgency of an item. An item's urgency depends on the frequency at which it is expected to be answered by users and the time that has passed since a user has last answered it. Our evaluation has revealed that there are questions for which a high repetition frequency is reasonable (such as "At the moment I have the feeling that I can cope with the requirements of my daily work") while others refer to phenomena which

are less dynamic, e.g. someone's identification with the company's goals and intentions.

The redesign implemented urgencies that range from zero to six; the higher the value, the more urgently an item needs to be answered. Zero and six mark special cases and can be used to indicate inactive items (0) and items that need to be presented to the user immediately (6). An item's urgency is always determined in relativity to its designated frequency. This means, for example, that if an item that should be answered every day has not been answered for a week, it has a higher urgency than an item that should be answered only once a month, even though it has also not been answered for one week. A lookup-table for the determination of item urgencies follows that includes two columns of examples with items of specific frequencies. Instead of the frequency f it uses its multiplicative inverse, the period $T = f^{-1}$.

urgency	generic	item with $T = 1$ week	item with $T = 1$ month
1	$0 < t \leq 0.5T$	$t \approx 0$ to 3 days	$t = 0$ to 2 weeks
2	$0.5T < t \leq T$	$t \approx 3$ to 7 days	$t = 2$ weeks to 1 month
3	$T < t \leq 1.5T$	$t \approx 7$ to 10 days	$t = 1$ month to 6 weeks
4	$1.5T < t \leq 2T$	$t \approx 10$ to 14 days	$t = 6$ weeks to 2 months
5	$t > 2T$	$t > 2$ weeks	$t > 2$ months

Table 2. Item urgencies; calculation and examples (t = time that has passed since last answering of item; T = designated item period).

We have observed that users sometimes do not find time or interest to answer questions within a longer period of time. This can result in many items of the questionnaire reaching a maximum urgency value (when the time since the last answer is $> 2T$) and thus become relevant. This would eliminate the desired effect of partitioning the questionnaire with the help of the relevancy concept. To respect timeouts from participation while avoiding such an agglutination, an item's urgency is reset to one, when t reaches $2.5T$. Figure 6 depicts how an item's urgency changes over time in case it is not answered (which would reset its urgency to 1 earlier).

The user's acceptance threshold. The counterpart to the item urgency is the user's acceptance threshold. We have observed that there are significant inter- and intra-personal differences in the perception of an acceptable participation quota (i.e. number of items per time). Consequently, we have introduced the acceptance threshold. It represents the user's current willingness to participate in the survey by answering items. If an item's urgency is of a higher value than a user's threshold, the item is considered relevant and

should be presented to the user. If a user had a static acceptance threshold of two, the item would become relevant again at exactly the designated time period T after the item had last been answered.

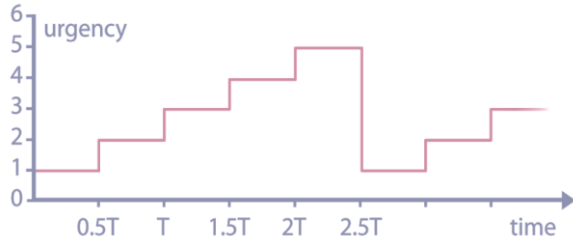


Figure 7. An item's urgency increases relatively to the item's frequency with time.

In general, it can be said that the higher a user's acceptance threshold, the lower the frequency is at which items become available to her or him (see Figure 8).

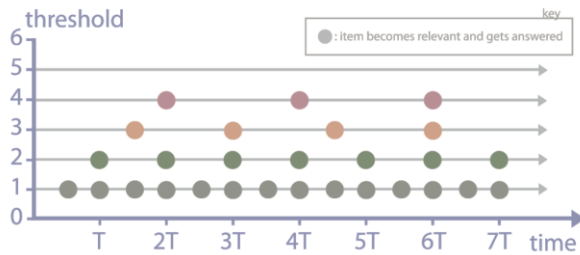


Figure 8. The frequencies at which items become relevant depends on the user's acceptance threshold.

While applying different acceptance thresholds to users is enough to pay regard to interpersonal differences in participation willingness, the system also supports changes on the intrapersonal level. The qualitative evaluation showed that the participation quota that is perceived as acceptable by users changes over time, as it is effected by external influences like workload. This means that the number of items which the survey system wants its participants to answer can change from person to person, but also for one person over time. To reflect this, a mechanism to determine a user's current acceptance threshold is needed. The redesign uses a combination of adaptability and adaptivity of the threshold's level:

Adaptability of the acceptance threshold. As it is shown in Figure 8, the acceptance threshold influences the frequencies at which items are relevant for the user and thus the number of items that are presented to the user. The system allows users to configure their own acceptance threshold by setting its absolute value directly or by giving feedback about the current number of items, which may lead to a relative

change of the threshold's value (e.g. stating 'there are too many questions' makes the system increment the acceptance threshold). In this way, the user determines the threshold's base value, which represents the user's general willingness to participate over several days or weeks.

Adaptivity of the acceptance threshold. Literature research showed that it is a highly complex task to use context information to determine if a situation is suitable to prompt a user without risking a disruption and in turn, annoyance (especially in real-world scenarios, like surveying during daily work). Thus it was decided to not solely rely on context awareness, but to instrumentalize it as a supplementary mechanism that could influence the acceptance threshold. If a contextually aware component of the system detects a situation that is favorable for asking the user to answer an item, it can apply a negative offset to the threshold. This could result in items becoming relevant and thus being presented to the user. When the detected situation has passed, the threshold returns to its original value. This process is depicted in Figure 9. The contextual influence adds an offset to the threshold's base value; this offset lasts only minutes to hours.



Figure 9. Item relevancy changing due to a temporarily decreased acceptance threshold.

In summary: A user's acceptance threshold has a base value, which can be altered by feeding back opinions about the amount of items that the system currently expects the user to answer. Contextual information can be used to add short-situational positive or negative offsets to the acceptance threshold's base value. The next section describes how context information is put to use in the proposed survey system.

4.2 A Context Awareness Framework

Research on context aware systems has shown (see Section 2) that considering context awareness is a promising approach to detect the appropriate points in time at which users could be asked to answer survey items. But implementation of this approach has not

matured enough to master such a complex task automatically. Still it can be of help to minimize disruptions and thereby maintain user acceptance for the survey over time.

To integrate context awareness into the system, the acceptance threshold is used as an interface between the application and context awareness modules (CA-modules). The acceptance threshold is an abstraction of the user's current willingness to participate, which is simplified to an integer value from one to five. The CA-modules have the task of translating sensed context information into an offset for the acceptance threshold. The CA-framework we implemented for this task (see Figure 10) differs from conventional context aware systems in that the network layer (1) and part of the middleware layer (2a) are outsourced to multiple CA-modules. The modules sense context data and infer information that may be of relevance for the system. The proposed open middleware layer (2b) collects the input of all modules and, in combination with the threshold's base value, which is defined by user feedback, aggregates the information into the threshold's current value.

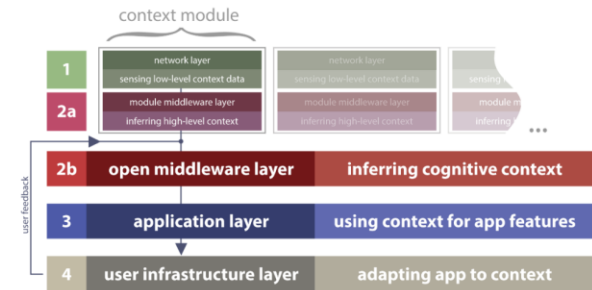


Figure 10. Modified layer architecture: Context inference is outsourced to CA- modules, which can access the system's open middleware layer.

Figure 11 gives an example of how contextual information could influence a user's current threshold. The shown scenario features three CA-modules, which are realizable on smartphone clients, as they use context data from common sensors and chips: A 'commuter module' checks the user's location and the current time and infers from this information that the user is probably commuting to work, which is assessed as a favorable situation for bothering the user with an item. To communicate this detection of a favorable situation, a negative offset (-1) is sent to the module that manages the acceptance threshold. A 'weekday module' checks for the date and compares it to the user's preferences. Because it 'knows' that the user does not want to participate at the weekend, the fact that it's Sunday induces the module to add a quite big positive offset (+3) to the threshold. The third module listens for ambient sounds, but as it cannot infer a situation of

relevance from the currently sensed data, it does not influence the threshold. The open middleware layer orchestrates the CA-modules' inputs and aggregates the thresholds current base value (4) with the offsets, which results in a threshold value of six. Based on this threshold value, the system can determine which subset of questionnaire items is relevant and may be presented to the user.

The main reason for decoupling context aware features from the system by outsourcing them to external modules is the current state of research in the field of context awareness. Literature research has shown that approaches to implementing context awareness are manifold and heterogeneous (see [28] for an overview and categorization of CA-systems). A determination of what is best practice in terms of abstraction, formalism, modelled dimensions or inference mechanisms is not possible. By offering the acceptance threshold as an interface for contextually influencing application behavior, the approach we followed was to delegate the task of deciding how to realize context awareness to the CA-modules. The internal inference mechanism of a context module is irrelevant for the system. This makes it easy to test different strategies for inferring results from contextual data and makes the system future-proof, as the CA-modules can be exchanged or extended without needing to alter application logic on the other end of the acceptance threshold.

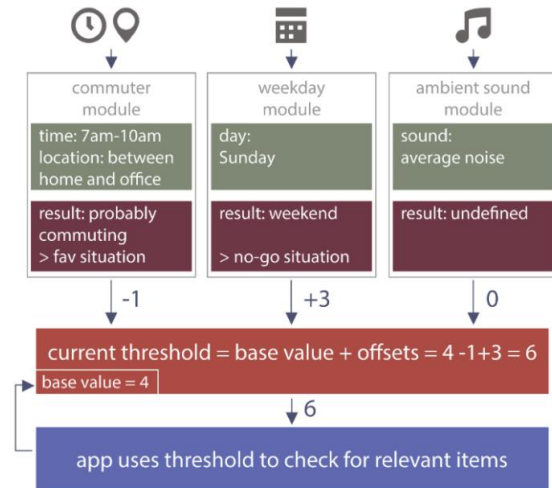


Figure 11. Example of the CA-modules' influence on a user's current threshold.

In addition to influencing the threshold, third-party applications might also profit from knowing the threshold's status. Developers could try to reduce users' notification load by orchestrating their general applications' notifications based on the acceptance threshold's status. However, the acceptance threshold

concept entails a potential risk, as granting access to the acceptance threshold could be exploited by developers, e.g. blocking other notifications by increasing the threshold while sending notifications themselves. A further advantage of modularizing context awareness lies in the oftentimes precarious nature of the data collected to infer knowledge about a user's situation. Whereas disclosing one's location to applications is becoming more and more common (e.g. for navigation purposes or for adding meta-data to posted status reports), users might have privacy concerns about giving an app access to the device's microphone or camera. Using modules allows users to make dedicated decisions about what privileges they grant to which module for which purposes. One trusted service could be accountable for inferring context data and computing cognitive context metadata (like the current acceptance threshold), while multiple, but probably less trustworthy apps, could make use of this information without accessing the device's sensors, like a camera or microphone.

Since agile companies promote the mobility of their employees, the presented redesign was prototypically implemented as a smartphone app for the Android operating system. Figure 12 presents its user interface, focusing on the functionalities that enable users to give relative feedback about the current number of questions they are asked. The left side of Figure 12 shows a system notification that indicates a survey question (A). The user can react by dismissing it (by swiping), increasing the acceptance threshold (by giving feedback that there are currently too many questions being asked – A2), assessing the contextual circumstances of the situation in which the notification appeared (e.g. that it's weekend – A3) or by accessing the item via the app (A1). When an item is loaded within the app (Figure 12, right), the user has more detailed possibilities to change her acceptance threshold. Clicking B1 or B2 increases the threshold by 2, respectively 1, while choosing B3 decreases the threshold by 1. These relative changes to the acceptance threshold are optional and can be ignored by simply answering relevant survey questions or closing the app.

5. Summary & Future Work

As discussed, continuous monitoring of team climates can be beneficial for fostering and maintaining working conditions that promote creativity and innovation, especially in agile corporate environments. This paper presents the design rationale and prototypic implementation of a survey system which supports the continuous data-collection that is needed for such monitoring. The prototype described here is a proof of an agile, socio-technical concept where participants of

a survey co-determine how it is run. Consequently, it differs from common survey systems that are not compatible with the flexibility requirements and the dynamics of creativity-oriented sites (see 3), or with the high priority of user acceptance. In particular, the concept features are:

- A specific repetition rate for every question item. This enables the designers of the questionnaire to adapt the item's frequency to the rate at which the question's subject is likely to change. A well-chosen frequency can increase the meaningfulness that is perceived by the participants when repetitively answering the item.
- A high self-determination of users about their rate of participation, which respects the intrapersonal and interpersonal differences in the users' willingness to participate.
- An open context awareness component, which can detect favorable and unfavorable situations for prompting the user with a question.

Our experiments revealed first insights into team climate dynamics: It appears that characteristics which influence creativity (e.g. stress or an appropriate workload) can be subject to significant change even within a few weeks. This observation supports the initial reasoning for continuous monitoring.

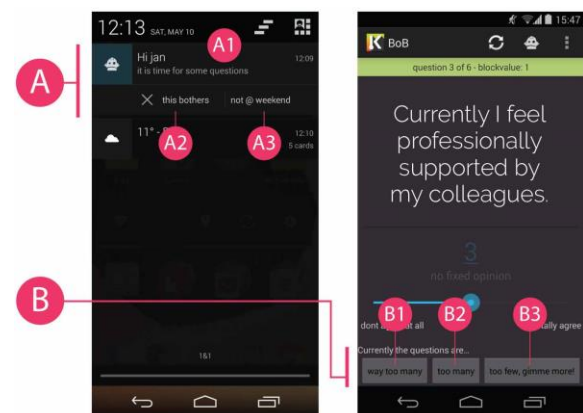


Figure 12. User interface of the Android survey client.

Future research needs to evaluate whether the proposed system design maintains user acceptance over a long period of time. The system can be used to learn more about the dynamics of creative work climates and their possible patterns, as well as the efficiency of interventions that are based on the monitoring. A consistent next step would be to feed the survey results back into the system, to automatically adapt the items' frequencies: the faster changes take place, the higher the monitoring frequency should be.

Another remaining research challenge is caused by the flexibility with which the questions can be answered.

This leads to heterogeneous data sets. Therefore, statistical methods have to be chosen and tested which allow for reliable and easily understandable feedback to the participants about the survey's results and the effects of changing work conditions on the creativity climate.

6. References

- [1] T. M. Amabile, R. Conti, H. Coon, J. Lazenby, and M. Herron, "Assessing the work environment for creativity," *Academy of Management Journal*, vol. 39, no. 5, pp. 1154–1184, 1996.
- [2] R. W. Woodman, J. E. Sawyer, and R. W. Griffin, "Toward a theory of organizational creativity," *Academy of management review*, vol. 18, no. 2, pp. 293–321, 1993.
- [3] C. M. Ford, "A theory of individual creative action in multiple social domains," *Academy of Management review*, vol. 21, no. 4, pp. 1112–1142, 1996.
- [4] T. M. Amabile and R. Conti, "Changes in the work environment for creativity during downsizing," *Academy of Management Journal*, pp. 630–640, 1999.
- [5] J. Nierhoff, "Cognitive Context Awareness - A Concept for Self-Determined Repetitive Data Collection within Surveys," 2016.
- [6] B. E. Ashforth, "Climate formation: Issues and extensions," *Academy of management review*, vol. 10, no. 4, pp. 837–847, 1985.
- [7] B. Schneider and A. E. Reichers, "On the etiology of climates," *Personnel psychology*, vol. 36, no. 1, pp. 19–39, 1983.
- [8] I. Borg and P. M. Mastrangelo, *Employee surveys in management: Theories, tools, and practical applications*. Hogrefe & Huber Publishers, 2008.
- [9] T. M. Amabile, P. Goldfarb, and S. C. Brackfield, "Social influences on creativity: Evaluation, coercion, and surveillance," *Creativity research journal*, vol. 3, no. 1, pp. 6–21, 1990.
- [10] K. Byron, S. Khazanchi, and D. Nazarian, "The relationship between stressors and creativity: a meta-analysis examining competing theoretical models," *Journal of Applied Psychology*, vol. 95, no. 1, p. 201, 2010.
- [11] E. L. Santanen, R. O. Briggs, and G.-J. D. Vreede, "Causal relationships in creative problem solving: comparing facilitation interventions for ideation," *Journal of Management Information Systems*, vol. 20, no. 4, pp. 167–198, 2004.
- [12] T. Herrmann and A. Nolte, "The integration of collaborative process modeling and electronic brainstorming in co-located meetings," in *International Conference on Collaboration and Technology*, 2010, pp. 145–160.
- [13] O. Neuberger, "Mitarbeiterbefragungen als symbolische Politik," *Mitarbeiterbefragung—Ein Instrument des Innovations-und Qualitätsmanagements*. Weinheim, pp. 423–434, 1997.
- [14] T. Herrmann, A. Carell, and J. Nierhoff, "Creativity barometer: an approach for continuing micro surveys to explore the dynamics of organization's creativity climates," in *Proceedings of the 8th ACM conference on Creativity and cognition*, 2011, pp. 345–346.
- [15] Allen I. Kraut, *Getting Action from Organizational Surveys: New Concepts, Technologies, and Applications*. Pfeiffer, 2006.
- [16] R. Hossiep and P. Friege, "Der Einsatz von Mitarbeiterbefragungen in Deutschland, Österreich und der Schweiz," *Planung und Analyse*, vol. 6, pp. 55–59, 2008.
- [17] M. E. Domsch and D. H. Ladwig, "Mitarbeiterbefragungen—Stand und Entwicklung," in *Handbuch Mitarbeiterbefragung*, Springer, 2013, pp. 11–55.
- [18] J. R. Evans and A. Mathur, "The value of online surveys," *Internet Research*, vol. 15, no. 2, pp. 195–219, 2005.
- [19] F. Reiss and S. Weltzin, "Mitarbeiter-und Führungskräftebefragungen aus datenschutzrechtlicher Perspektive, tivian-Whitepaper," *Verfügbar unter: <http://www.tivian.de/deutsch/download.html>*, 2010.
- [20] J. M. Hudson, J. Christensen, W. A. Kellogg, and T. Erickson, "I'd be overwhelmed, but it's just one more thing to do: Availability and interruption in research management," in *Proceedings of the SIGCHI Conference on Human factors in computing systems*, 2002, pp. 97–104.
- [21] S. B. Shneiderman and C. Plaisant, *Designing the user interface 4th edition*. Pearson Addison Wesley, USA, 2005.
- [22] L. Xia and D. Sudharshan, "Effects of interruptions on consumer online decision processes," *Journal of Consumer Psychology*, vol. 12, no. 3, pp. 265–280, 2002.
- [23] C. A. Monk, D. A. Boehm-Davis, G. Mason, and J. G. Trafton, "Recovering from interruptions: Implications for driver distraction research," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 46, no. 4, pp. 650–663, 2004.
- [24] D. D. Salvucci, N. A. Taatgen, and J. P. Borst, "Toward a unified theory of the multitasking continuum: from concurrent performance to task switching, interruption, and resumption," in *Proceedings of the 27th international conference on Human factors in computing systems*, New York, NY, USA, 2009, pp. 1819–1828.
- [25] J. G. Trafton, E. M. Altmann, D. P. Brock, and F. E. Mintz, "Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal," *International Journal of Human-Computer Studies*, vol. 58, no. 5, pp. 583–603, 2003.
- [26] P. D. Adamczyk and B. P. Bailey, "If not now, when?: the effects of interruption at different moments within task execution," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2004, pp. 271–278.
- [27] E. Horvitz, C. Kadie, T. Paek, and D. Hovel, "Models of attention in computing and communication: from principles to applications," *Communications of the ACM*, vol. 46, no. 3, pp. 52–59, 2003.
- [28] J. Hong, E. Suh, and S.-J. Kim, "Context-aware systems: A literature review and classification," *Expert Systems with Applications*, vol. 36, no. 4, pp. 8509–8522, 2009.
- [29] M. Csikszentmihalyi, "Flow and the psychology of discovery and invention," *New Yprk: Harper Collins*, 1996.
- [30] F. C. Brodbeck and G. W. Maier, "das Teamklima-Inventar (TKI) für Innovation in Gruppen," *Zeitschrift für Arbeits-und Organisationspsychologie*, vol. 45, no. 2, pp. 59–73, 2001.
- [31] J. Lapierre and V.-P. Giroux, "Creativity and work environment in a high-tech context," *Creativity and Innovation Management*, vol. 12, no. 1, pp. 11–23, 2003.